

Metulini, Rodolfo

Conference Paper

A structural analysis on Gravity of Trade regarding the possibility to remove distance from the model

53rd Congress of the European Regional Science Association: "Regional Integration: Europe, the Mediterranean and the World Economy", 27-31 August 2013, Palermo, Italy

Provided in Cooperation with:

European Regional Science Association (ERSA)

Suggested Citation: Metulini, Rodolfo (2013) : A structural analysis on Gravity of Trade regarding the possibility to remove distance from the model, 53rd Congress of the European Regional Science Association: "Regional Integration: Europe, the Mediterranean and the World Economy", 27-31 August 2013, Palermo, Italy, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:

<http://hdl.handle.net/10419/123952>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

A structural analysis on Gravity of Trade regarding the possibility to remove distance from the model

Rodolfo Metulini

March 1, 2013

Abstract

The Gravity Model is the workhorse for empirical studies in International Economies and it is commonly used in explaining the trade flow between countries. Recently, several studies have showed the importance of taking into account the spatial effect. The standard procedure consists on considering the transport costs using geographical distance as a proxy, and the spatial effect through a weighted matrix constructed on inverse distance. Two issues follow from this standard procedure: the first regards the biasness of the distance if it is used as a proxy of the transport costs, the second is related to the collinearity emerging if we use distance twice. So, several attempts were made in the recent literature having the scope of removing the distance. We propose a theoretically consistent procedure based on Anderson, Van Wincoop derivation model, and some ad-hoc tests, relating to this attempt. The empirical results based on a 22-years panel of OECD countries are comforting, and they allow us to estimate the model without the distance, if properly replaced by a set of fixed effects.

1 Introduction

The literature shows how, at empirical level, the classical Gravity Model (Anderson 1979, Anderson, Van Wincoop 2003) brings to good results in explaining international trade. Anderson (1979, p.106) states that this equation is the most successful in explaining this issue, furthermore, Everett and Hutchinson (2002, p-489) define this model as the workhorse for empirical studies in international economy. Since the gravity model has physical roots, the trade flow depends on the economic dimension of both the origin and the destination countries, as well as the trade costs between them.

From the beginning of the gravity model of trade, trade costs were taken into account using the distance variable and other variables considered relevant on the volume of trade, such as geographic contiguity, common language and common currency, and the presence of some Free Trade Agreements.

However, the original model by Anderson did not consider some effects related to the (auto)correlation in space. Recently, several studies highlighted that an additional effect to take into consideration is the spatial dependence. Following Tobler's first law of Geography (1970): *everything is related to everything else, but closer things are more closely related than distant things*, regional scientists began to give relevance to this aspect.

Spatial dependence can be explained by the *third country* theory or by the presence of some spatially correlated omitted variables. The Econometric literature suggests to model this phenomenon with appropriate spatial autoregressive models and methods.

Moreover, the unconstrained model derived by Anderson, Van Wincoop (2003) theorized the presence of the *multilateral resistance terms*, that accounts for the openness of each partner and reporter country. They are usually proxied by fixed effects.

The usual procedure in the definition of the formulation of the spatial gravity model, which is becoming a gold standard, is to account for the trade cost using geographical distance and other dummy variables (contiguity, language, currency) as a proxy and the spatial effect through a weighted matrix constructed on inverse distance. Two issues follow from this standard procedure: the first is related to the collinearity emerging if we use distance twice, that produce larger standard error of the estimated coefficients; the second regards the biasness of the distance if used as a proxy of the trade costs. This biasness emerges for several reasons: it is related to 1) the changing nature of the trade costs that distance cannot account for¹, 2) the unsuitable choice of the centroid used to define distance, 3) the fact that different kind of transport can have different costs. This last fact is widely argued in the literature (Head, Mayer 2002, Martinez-Zarzoso, Suarez-Burguet 2005, Anderson, Van Wincoop 2004, Cheng, Wall 2005).

The above motivations justify us to test for the possibility to remove the distance variable from the gravity model, which is the goal of this work.

Cheng and Wall adopted for this scope a nested approach using Likelihood ratio (LR) tests and they find that fixed effects are suitable to use instead of distance. A theoretically consistent procedure (which is despite devoted to a little bit different goal) is indeed proposed by Anderson, Yotov (2012).

Since, as we already said, the MRT's are usually proxied by fixed effects, we are going to check on the possibility to proxy the *multilateral resistance terms* and the *transport costs* (standardly proxied by distance) jointly with a *set* of fixed effects, motivating that the most accurate choice for fixed effects lies in the couple of country specification without symmetry constraint.

On the light of the Anderson, Yotov (2012) procedure, our proposal is about determining a *multilateral resistance terms* estimated coefficients, through the derivation of the structural model derived by Anderson, Van Wincoop (2003), jointly with an estimation of the empirical fixed effects from an empirical model

¹trade costs change over time

which doesn't include distance as a proxy for the transport costs. In other words, these coefficients will be determined so that the fixed effects account for the MRT's and the distance effects jointly. Therefore, they represent a suitable proxy for MRT's and distance jointly. Under such hypothesis, distance and *multilateral resistance terms* should be equal to the fixed effects unless for a random effect distributed as a white noise (i.e. normal with zero mean).

We will perform an ANOVA analysis to study for the MRT's plus distance variance explained by fixed effects. Moreover, a bootstrap analysis to test for the normal distribution with zero mean of the residuals calculated as the difference between MRT's plus distance and the fixed effects.

The results of the just mentioned tests confirm our hypothesis that we can remove the distance if we use a suitable set of fixed effects.

In the next paragraph we discuss the motivation underlying this work. It follows a theoretical section in which we show the framework and how to derive the estimated structural components object of comparison for our analysis. Then, in the fourth section, we show some ad-hoc suitable test and in the fifth section we present an empirical analysis based on a panel of OECD countries in order to check for the above discussed theories. The last section conclusions will be discussed.

2 Motivations

The usual procedure in the definition of the right formulation of the spatial gravity model, which is becoming a gold standard, is to account the trade costs using geographical distance and other dummies (such as contiguity, language, currency) as a proxy, and the spatial effect through a weighted matrix constructed on inverse distance.

Two issues follow from this standard procedure: the first regards the biasness of the distance if used as a proxy of the transport costs, the second is related to the collinearity emerging if we use distance twice.

Starting from the second issue, we said about the adoption in the spatial model of the contiguity matrix W constructed through the inverse of the distance. We are going to have the distance variable used twice in the model. Despite this is, as we already said, a standard practice in the literature, evidences of collinearity aspects are likely to be among the explanatory variables. Collinearity among explanatories determines an increase of the standard errors of the estimated coefficients.

Relating to the first issue, it is widely argued in the literature (Head, Mayer 2002, Martinez-Zarzoso, Suarez-Burguet 2005, Anderson, Van Wincoop 2004) that distance variable is not a suitable proxy in accounting for the *bilateral resistance* effects that the theoretical gravity model identify with trade costs (Anderson, Van Wincoop, 2003). Martinez-Zarzoso and Suarez - Burguet (2005) highlighted that trade costs would be inversely correlated with commercial flow and should be modeled together in a two equations system. This changeable nature of the trade costs make the distance variable inappropriate to proxy the

bilateral resistance, since it doesn't change over time.

Moreover, using distance we suffer of missmeasurement, since transport costs could change along different type of transport. For example the cost to overland goods by sea is much cheaper than overland goods by land (Cheng, Wall, 2005). Distance is not able to account for this fact.

Another aspect stressed by Cheng and Wall is related to the choice of the centroids in measure the distance between the two countries: sometimes, the choice of the point in space in which we choose the centroid do not fit with the economic center of the countries (in which the main part of the commercial flows happens). Furthermore, it is empirically demonstrated that the introduction of the distance variable in the model overestimates the negative effect of the *bilateral resistance* (Rose 2002, Cheng, Wall 2005).

One solution could be to use the transport costs to account for itself: in this case the problem disappear. Nevertheless, data on transport costs are hard to obtain: this is the reason why usually we account this by distance proxy.

So, the above motivations justify us to test for the possibility to remove the distance variable from the gravity model. Cheng and Wall adopted for this scope a nested approach using Likelihood ratio (LR) tests. They compare a model with distance and a model with fixed effects. This approach does not have solid economic foundations. A theoretically consistent procedure (which is despite devoted to a little bit different goal) is indeed proposed by Anderson, Yotov (2012), which is based on the economic derivation of the structural gravity model by Anderson, Van Wincoop (2003)

3 Theoretical procedure to derive the estimated structural terms

To derive the estimated coefficients we spoke about before, we make use of the derivation of the structural model by Anderson, Van Wincoop (2003).

To derive this model we adopt the concept of trade separability: for this concept, at an upper level we generate the production and the expenditure for each good in each region, and at a lower level the demand and supply, conditioned to the production and expenditure value determined before.

To derive the model we make use of the following constraints:

- *budget constraint* (one for each destination);
- *market clearance condition* (one for each origin);

jointly with the use of the *CES* demand function.

Defining $T_{ij} \geq 1$ as the variable relative to the whole trade costs and σ the elasticity of substitution, we define the *CES* demand function.

$$Y_{ij} = (\beta_i p_i T_{ij} / P_j)^{1-\sigma} E_j \tag{1}$$

The budget constraint determines on the relative price:

$$P_j = [\sum_i \beta_i p_i T_{ij}]^{1-\sigma}]^{1/1-\sigma} \quad (2)$$

Assuming the *market clearance condition*, in other words $X_i = \sum_j X_{ij}$, we sum over j the *CES* demand function to obtain that:

$$X_i = \sum_j (\beta_i p_i)^{1-\sigma} (T_{ij}/P_j)^{1-\sigma} E_j. \quad (3)$$

Using $X = \sum_i X_i$, we obtain that

$$X_i/X = (\beta_i p_i \Pi_i)^{1-\sigma} \quad (4)$$

where $\Pi_i = \sum_j (T_{ij} P_j)^{1-\sigma} E_j / X$

For the *market clearance condition* constraint, if we sum over i the last equation we obtain that $\sum_i (\beta_i p_i \Pi_i)^{1-\sigma} = 1$.

Using the same equation to substitute his terms p_i e β_i in the initial demand function, to obtain a three equation system representing the gravity model:

$$Y_{ij} = \frac{E_j X_i}{X} \left(\frac{T_{ij}}{P_j \Pi_i} \right)^{1-\sigma} \quad (5)$$

$$(\Pi_i)^{1-\sigma} = \sum_j \left(\frac{T_{ij}}{P_j} \right)^{1-\sigma} \left(\frac{E_j}{X} \right) \quad (6)$$

$$(P_j)^{1-\sigma} = \sum_i \left(\frac{T_{ij}}{\Pi_i} \right)^{1-\sigma} \left(\frac{X_i}{X} \right) \quad (7)$$

Hence, the gravity model derived by Anderson and Van Wincoop considers the *multilateral resistance effects*, which are unknown, but we can estimate them through a three equation system defined above.

The equations (5) – (7) are fondant for our analysis: their stochastic versions permit us to estimate the theoretical *multilateral resistance effects*. These estimates are based on a previous estimate of the E_j , X_i and T_{ij} components, where each of these will be defined on the basis of empirical values, as we will see.

We can estimate the empirical values trough the equation (8).

$$Export_{ij} = (\hat{X}_i) + (\hat{E}_j) + (\hat{T}2_{ij}) + (others_{ij}) + \theta_{ij} + \varepsilon_{ij} \quad (8)$$

where θ_{ij} should account for the effect of the transport costs and of the *multilateral resistance terms*. Moreover, with \hat{X}_i we refer to an estimated value for the reporter economic dimension, with \hat{E}_j we refer to an estimated value for the partner economic dimension, and \hat{T}^2_{ij} represent the estimated trade costs different from transport costs (total trade costs – transport costs)

This step, in fact, is twofold: we estimate with the same equation model: a) a set of $n * n$ fixed effects defined such that they account for the effect of the *multilateral resistance* and the effect of transport costs, b) the estimated structural terms of the gravity.

The choice of the right set of fixed effects is quite relevant: the distance accounts for the transport costs between couple of country, which could be different along different directions (from i to j or from j to i). So, to use a set of symmetric effects could not be suitable. At the same, to consider two sets of fixed effects, one for the *reporter* country, the other for the *partner* country, could be worse, since couple fixed effects account for much more variability.

One aspect of debate regards the choice to identify the distance as the only proxy variables of trade costs we want to remove. That is because we assume the other proxies for trade costs:

- unbiased and without identification problem;
- of standard use in the model.

Hence, assuming this hypothesis, we need for an evidence of the decomposability of trade costs. Feenstra (1998) shows that the total trade costs are due for the 21% to the transport costs (for which distance is used as a proxy). The other part of the cost are motivated by the market entry barriers.

We can assume $\hat{T}_{ij} = dist_{ij} + \hat{T}^2_{ij}$ be the equation defining the total trade costs decomposition in transport costs and other *bilateral resistance costs* (T^2).

To resume, we determine by a stochastic version of the equations (5) – (7), with support of (8), the estimated coefficients for *multilateral resistance terms* $[\hat{\Pi}_i; \hat{P}_j]$, and the estimated set of couple fixed effect.

4 Proposed analysis and tests

The structural model theory would implies that the expected value of the estimated fixed effect would be equal to the expected value of the estimated *multilateral resistance terms* plus distance. Differences might only be due to random deviation around the zero (i.e. the difference is distributed as a white noise):

$$\theta_{ij} = \Pi_i + P_j + dist_{ij}.^2 \tag{9}$$

²Here, the fact that we do not have hat indicates that we think this relation at an expected value level

It is possible to test the above mentioned equation by comparing those estimated components:

$$r_{ij} = \hat{\theta}_{ij} - (\hat{\Pi}_i + \hat{P}_j + dist_{ij}). \quad (10)$$

The r_{ij} term could be interpreted as a residual component of a regression model of the estimated fixed effects $\hat{\theta}_{ij}$ respect to $\hat{\Pi}_i + \hat{P}_j + dist_{ij}$, with the constraint that the intercept must be equal to 0 and every slope parameter is equal to 1 .

That interpretation permit us to analyze the r_{ij} component through an analysis of variance (*ANOVA*, as made by Egger, Pfaffermayr (2002, 2004), as a fitting measure for the structural components. The ANOVA is considered as a special case of the linear regression model (Gelman 2005, Montgomery 2001). Both, indeed, consider the observed θ_{ij} as a summation of the structural variables and the residuals.

The statistic significance of the *test* is determined by the ratio between the two variance. In other words, placing:

- $V(r)$ the variance of r_{ij} ,
- $V(S)$ the variance of $\hat{\Pi}_i + \hat{P}_j + dist_{ij}$;

the proportion of the variance of the fixed effects $\hat{\theta}_{ij}$ explained by the components of the structural model $\hat{\pi}_i + \hat{P}_j + \widehat{dist}_{ij}$ will be $\frac{1-V(r)}{V(S)}$, which is the correlation index R^2 of the regression model if we assume the above constraints.

To clarify, with this test we want to evaluate how much variability of the fixed effects is explained by their theoretical counterpart (i.e. the MRT's used on the structural model and the distance as a proxy of transport costs).

Moreover, a further hypotesis must be assumed: the hypotesis of indipendence of test from possible alteration of the experiment. This hypotesis is satisfied, because both the dependent and the explanatories in the *ANOVA* test were defined conditionally to the same structural variables composing the model.

The interest for the validation of the empirical model without distance need for a further suitable test for the null hypothesis regarding the expected value of the r_{ij} 's.

Unfortunately, an hypotesis test like the *t - test* to evaluate the null $H_0 : E(r_{ij}) = 0$ needs the assumption that such residuals are generated from an independent and identically distributed (i.i.d.) process, and that the sample mean of the r_{ij} 's is normally distributed. The second assumption is ensured from the result of the central limit theory.

Instead, the i.i.d. assumption is hardly conceivable, in fact, such residuals are generated from a system that jointly considers the *multilateral resistance terms* (Π_i e P_j) and the distance ($dist_{ij}$), which components are each other correlated: in fact, the *multilateral resistance terms* are estimated through the equations (6) and (7) using the distance as true value of the transport cost among the regressor, which contains a measurement error (since it is a proxy).

Such measurement error affects the estimated *multilateral resistance terms*, that will be correlated with the distance.

A possible solution could be to perform a *t – test* based on *bootstrap* procedure, bootstrapping directly the r_{ij} residuals. This procedure presents some problems, in fact, the bootstrap procedure need for the i.i.d. assumption of the bootstrapped residuals. Since we are going to have dependent residuals, a solution proposed in the literature could be to perform the *block – bootstrap*. However, this approach need to vary the dimension of the sub-samples (see Politis, Romano, 1994).

Conversely, is possible to assume that the ε_{ij} residuals from model in (8) are i.i.d.. Through a *regression bootstrap* procedure which use the residuals in (8) as a starting point, is possible to obtain a set of r_{ij} *bootstrapped* residuals which are i.i.d..

Here, we describe the step by step procedure.

1. To estimate (8) and to obtain the estimated ε_{ij} ;
2. To perform a *bootstrap* on the residuals generated at the first step: we are going to obtain $[\varepsilon_{ij,1}, \dots, \varepsilon_{ij,B}]$. We compute a number equal to B *bootstrapped* terms for $Export_{ij}$:

($[Export_{ij,1}, \dots, Export_{ij,B}]$) through the following relation:

$$Export_{ij,b} = \widehat{Export}_{ij} + \varepsilon_{ij,b}; \forall b = 1, \dots, B$$

where \widehat{Export}_{ij} is the estimation from the model (8);

3. To perform a number of B linear regressions of the model (8), using, for each one, each of the B *bootstrapped* terms for $Export_{ij}$ obtained at the second step; in order to obtain B sets of coefficients for the model (8) and B estimations for the fixed effects θ_{ij} ;
4. For each of the B iterations, to re-estimate the *multilateral resistance terms* throughout the relations defined in (6) and (7).

At the end of this procedure, we would have B *bootstrapped* estimates for the fixed effects: $[\theta_{ij,1}, \dots, \theta_{ij,B}]$. B *bootstrapped* estimates for the *multilateral resistance terms*: $[\Pi_{i,1}, \dots, \Pi_{i,B}]$, $[P_{j,1}, \dots, P_{j,B}]$. B estimates for the distance: $[dist_{ij,1}, \dots, dist_{ij,B}]$. To the end, we can define B *bootstrapped* i.i.d. terms representing the r_{ij} 's, so defined:

$$r_{ij,b} = \theta_{ij,b} - (\Pi_{i,b} + P_{j,b} + dist_{ij,b}); \forall b = 1, \dots, B.$$

Now, to compute a *t – test* become possible. For each iteration we compute the mean:

$$M(r_{ij,b}) = \frac{\sum_{ij} r_{ij,b}}{n^2},$$

And its *standard error*:

$$SE(r_{ij,b}) = \left[\frac{\sum_{b=1}^B M(r_{ij,b}) - M(r_{ij})}{\sqrt{n^2}} \right]^{-1/2},$$

where $M(r_{ij}) = \sum_{i=1}^B M(r_{ij,b})$ represents the mean of the means.

The *t-test* to evaluate the $H_0 : r_{ij} = 0$ hypothesis assumes the following form:

$$t \sim \frac{M(r_{ij})}{SE(r_{ij,b})} \quad (11)$$

5 Empirical Analysis

5.1 Choice of the empirical model

To test the proposed analyzes we use a panel related to the country-to-country trade flow over the period from 1988 to 2009 for the OECD members. The sample contains 32 countries, and the time series is 22 years long, resulting in a $n * n * T = 22528$ observations.

As the explanatory variables of the empirical gravity model we use the consumer price index at purchase power of parity (ppp), the gross domestic product (gdp) in real terms and the population (pop) for the origin and for the destination country. We use distance among centroids to proxy for the transport costs. Moreover, besides the dummies *contig*, *comcur*, and *comlang*, we use the dummies relatives to the free trade agreement of EU 15, NAFTA and EFTA. Furthermore, the dummy variables representing the effect of each couple of countries are inserted in the model and the stock of immigrants. The purchase power parity index is available at PennTable web site, in a dataset called PWT 7.0. The explanatory variable, as well, are available at the database of PennTable (gdp and pop), and from CEPII (*comlang*, *contig* and *comcur*).

Along with the spatial effects components, we define a weight matrix using the inverse of the distance between the centroids (available at CEPII as well).

As proposed by Arbia (2009) we check the spatial effects assumption by mean of the Moran I on the OLS residuals with the inverse distance matrix. This test show a significant spatial effects. So, by adopting a General to Specific approach, we assume as the true model the SARAR in which both the spatial lag and the error lag is taken into account. In order to choice between random versus fixed effects, a spatial Hausmann test was performed. This test highlights the strong preference for the fixed effects model. Furthermore, a Spatial Lagrange Multiplier test was performed to check the presence of spatial effects and spatial autocorrelation.

Since we have a spatial autoregressive model, we need to control for the endogeneity of the explanatories. We do so by using IV/GMM estimation method.

Therefore, based both on our tests and on the literature, we have reasons to choose a spatial model where distance appears twice (as proxy for transport costs and to define the weighted matrix): this follows the gold standard of the gravity models, where the spatial effects is taken into account through a weighted matrix constructed on inverse distance.

The empirical model takes therefore the following form:

$$\begin{aligned}
\text{Export}_{ijt} &= \alpha_{ij} + \alpha_t + \beta_1^o \text{Pop}_{it}^o + \beta_2^o \text{GDP}_{it}^o + \beta_3^o \text{ppp}_{it}^o + \beta_4^o \text{Nafta}_{it}^o + \beta_5^o \text{Efta}_{it}^o + \\
&\beta_6^o \text{Eu15}_{it}^o + \beta_1^d \text{Pop}_{jt}^d + \beta_2^d \text{GDP}_{jt}^d + \beta_3^d \text{ppp}_{jt}^d + \beta_4^d \text{Nafta}_{jt}^d + \beta_5^d \text{Efta}_{jt}^d + \beta_6^d \text{Eu15}_{jt}^d + \\
&\beta_7 \text{Migrat}_{ijt} + \psi_1^{od} \text{Contig}_{ij} + \psi_2^{od} \text{Comlang}_{ij} + \psi_3^{od} \text{Comcur}_{ij} + \psi_4^{od} \text{Dist}_{ij} + \\
&\psi_5 w_{ij,hk,OD} \text{bilat.PIL}_{ijt} + \rho w_{ij,hk,OD} y_{hk,t} + u_{ijt}; \\
u_{ijt} &= \lambda \sum_{hk=1}^{n^2} w_{ij,hk} u_{hk,t} + \varepsilon_{ijt}; \quad (12)
\end{aligned}$$

5.2 analysis and results

We can apply the analyzes proposed in in the previous section at theoretical level. First of all, we need to define the empirical model without distance specified in (8): we only removed the variable *dist* from (12):

$$\begin{aligned}
\text{Export}_{ijt} &= \alpha_t + \beta_1^o \text{Pop}_{it}^o + \beta_2^o \text{GDP}_{it}^o + \beta_3^o \text{ppp}_{it}^o + \beta_4^o \text{Nafta}_{it}^o + \beta_5^o \text{Efta}_{it}^o + \\
&\beta_6^o \text{Eu15}_{it}^o + \beta_1^d \text{Pop}_{jt}^d + \beta_2^d \text{GDP}_{jt}^d + \beta_3^d \text{ppp}_{jt}^d + \beta_4^d \text{Nafta}_{jt}^d + \beta_5^d \text{Efta}_{jt}^d + \beta_6^d \text{Eu15}_{jt}^d + \\
&\beta_7 \text{Migrat}_{ijt} + \psi_1^{od} \text{Contig}_{ij} + \psi_2^{od} \text{Comlang}_{ij} + \psi_3^{od} \text{Comcur}_{ij} + \alpha_{ij} \\
&\psi_5 w_{ij,hk,OD} \text{bilat.PIL}_{ijt} + \rho w_{ij,hk,OD} y_{hk,t} + u_{ijt}; \\
u_{ijt} &= \lambda \sum_{hk=1}^{n^2} w_{ij,hk} u_{hk,t} + \varepsilon_{ijt}; \quad (13)
\end{aligned}$$

In this equation, we've got the variable we must use to replace the terms of economic dimensions for partner and reporter countries, in order to obtain the estimated *MRT*'s via stochastic version of (6) and (7):

- $\hat{X}_{it} = \beta_1^o \text{Pop}_{it}^o + \beta_2^o \text{GDP}_{it}^o + \beta_3^o \text{ppp}_{it}^o$

- $\hat{E}_{jt} = \beta_1^d Pop_{jt}^d + \beta_2^d GDP_{jt}^d + \beta_3^d PPP_{jt}^d$

Moreover, the estimation for the trade costs T_{ij} comes from this proxy:

- $\hat{T}_{ij} = \psi_1^{od} Contig_{ij} + \psi_2^{od} Comlang_{ij} + \psi_3^{od} Comcur_{ij} + \psi_4^{od} Dist_{ij}$

We have to note that, with this application, we manage with panel data and with the t index for the time series. This permits us to obtain the estimation of the $n * n$ fixed effects, which is not possible for cross section, since the sample dimension is $n*n$. However, engaging a panel data analysis, we need to introduce some implementation to the theory showed in the previous section, in order to identify the estimated MRT 's. It is straightforward to implement the proposed methods in the previous sections for the case of panel data by adding summation over t in equations (6) and (7).

Proceed to an ANOVA analysis as described in the theory, with the aim to investigate how much of the fixed effects is explained by the structural variables.

The result is a R^2 of 0.9998.

This value, surprisingly high, tells us that the 99.98% of the total variance of the θ 's is explained by the structural model theorized by Anderson and Van Wincoop (2003), and gives us some demonstration of the fact that:

1. the structural model theorized by Anderson and Van Wincoop is a consistent formulation for the international trade phenomenon;
2. the effect of the distance can be accounted by a set of couple fixed effects;

Proceed to a Bootstrap analysis as described in the theory, now with the purpose to test if the expected value of r_{ij} is equal to zero.

The result of the $t - test$ showed in equation (11) assumes value equal to 1.12, which have a corresponding $p - value$ of $Pr(|T| > |t|) = 0.130$, therefore, the null hypothesis is not rejected. This analysis confirms the thesis for which the empirical gravity model can be consistently estimated without the inclusion of the distance variable, if properly replaced by couple fixed effects.

5.3 comparison and remarks

At an explorative level, to investigate the consistency of the previous analyzes, we are interested to compare the estimated coefficients of the two *SARAR* models, the first one with distance, the second without distance. Remember that the aim was to remove the distance from the model and the tests confirmed that this is possible. What happen to the estimated coefficients, once we remove distance variable from the model? We note that several variables increase their coefficients once the distance variable do not take part in the model (as we can see in table 1): the effects of the distance is therefore caught by the other variables. Nevertheless, the signs of the coefficients of the model without distance are in line with the literature.

Table 1: A preliminary comparison: *IV/GMM* estimated coefficients for SARAR models with and without distance

VARIABLE	without distance (*)	with distance
α	-	-
Contig	1.297 ($< 2.2e - 16$)	0.395 ($< 2.2e - 16$)
Comlang	0.877 ($< 2.2e - 16$)	1.084 ($< 2.2e - 16$)
Dist	-	-0.682 ($< 2.2e - 16$)
Comcur	3.335 ($< 2.2e - 16$)	0.549 ($< 2.2e - 16$)
Pop ^d	1.509 ($< 2.2e - 16$)	0.937 ($< 2.2e - 16$)
ppp ^d	0.289 ($< 2.2e - 16$)	0.095 ($< 2.2e - 16$)
GDP ^d	2.679 ($< 2.2e - 16$)	1.212 ($< 2.2e - 16$)
Pop ^o	1.793 ($< 2.2e - 16$)	0.953 ($< 2.2e - 16$)
ppp ^o	0.445 ($< 2.2e - 16$)	0.159 ($< 2.2e - 16$)
GDP ^o	2.484 ($< 2.2e - 16$)	1.170 ($< 2.2e - 16$)
Efta ^o	2.034 ($< 2.2e - 16$)	1.9269 ($< 2.2e - 16$)
Efta ^d	0.312 ($< 2.2e - 16$)	0.3981 ($< 2.2e - 16$)
Nafta ^o	2.743 ($< 2.2e - 16$)	-2.533 ($< 2.2e - 16$)
Nafta ^d	2.117 ($< 2.2e - 16$)	-1.993 ($< 2.2e - 16$)
EU15 ^o	0.830 ($< 2.2e - 16$)	0.7937 ($< 2.2e - 16$)
EU15 ^d	0.773 ($< 2.2e - 16$)	0.7236 ($< 2.2e - 16$)
W Bilat. PIL	-0.108 ($< 2.2e - 16$)	-0.080 ($< 2.2e - 16$)
Migrat	0.019 ($< 2.2e - 16$)	0.032 ($< 2.2e - 16$)
Rho	0.00055 ($< 2.2e - 16$)	0.00035 ($< 2.2e - 16$)
Lambda	($< 2.2e - 16$)-	0.00230 ($< 2.2e - 16$)

* In brackets the p-value

6 Conclusions and future developments

As previously motivated, the use of the *distance* variable have a critical role when we work with gravity of trade. Recent literature address this issue replacing with the fixed effects the distance used as a proxy for the transport costs. It was motivated the reason why the couple fixed effects without the hypothesis of symmetry are preferred to the couple fixed effects with hypothesis of symmetry; it was also motivated why fixed effects for reporter and partner countries are not suitable to proxy the transport costs. The analysis carried out here is based on the theoretical definition of the unconstrained gravity model, and it is consistent from an economic point of view. The analysis is based on a comparison between the estimated fixed fixed and the distance (plus the multilateral resistance terms estimated from the theoretical model).

The analyzes show interesting results: the first is that the structural model derived by Anderson, Van Wincoop seems to be consistent, since the theoretical *MRT*'s explain almost the total fixed effects variance. The second regards the confirmed possibility to remove the distance from the model: the effect of the *MRT*'s and of the transport cost is excellently grasped by a set of couple non-symmetric fixed effects.

The panel data application is very important: it allows us to use and estimate a set of n^2 fixed effects. This kind of fixed effects permit us to consider (and to proxy) more carefully the impact of transport costs, because we have one cost for each couple of countries, and they are not symmetric (costs from i to $j \neq$ costs from j to i).

A further development could be to investigate if at a sub-sectorial level this results still hold.

7 Bibliography

- Adam C., Cobham D. (2007): Modelling multilateral trade resistance in a gravity model with exchange rate regimes. Centre for dynamic macroeconomic analysis conference papers 2007.
- Anderson J.A. (1979): A Theoretical Foundation for the Gravity Model, The American Economic Review, 69,106-116.
- Anderson J.A., Van Wincoop E. (2003): Gravity with Gravitas: A solution to the Border Puzzleâ American Economic Review 93, 170-192.
- Anderson J.A., Van Wincoop E. (2004): Trade Costs. Journal of economic literature, 42, 691-751.
- Anderson J.A., Yotov Y.V. (2008): The changing incidence of geography. American Economic Review, vol. 100(5), p.2157-86.

- Anderson J.A., Yotov Y.V. (2010): Specialization: Pro and Anti-globalization, 1990-2002. NBER Working Paper No. 16301.
- Anderson J.A., Yotov (2012): Gold Standard Gravity, NBER Working Papers 17835, National Bureau of Economic Research, Inc.
- Anselin L. (1988): Spatial econometrics: Methods and Models (Kluwer Academic Publishers, Dordrecht)
- Baldwin R., Taglioni D. (2006): Gravity for Dummies and Dummies for gravity equation - NBER Working Paper No. 12516.
- Baltagi BH (2008): Econometric analysis of Panel Data, 4th edition, John Wiley e sons. Ltd.
- Egger P., Pfaffermayr M. (2002-1): The prime effect of the european integration intra EU core and periphery trade. Working paper in econometrics, University of Innsbruck.
- Egger P., Pfaffermayr M. (2002-2): Foreign direct investment and european interaction in the 90's. Working paper in econometrics, University of Innsbruck.
- Egger P., Pfaffermayr M. (2004): The impact of bilateral investment treaties on foreign direct investment, Journal of Comparative Economics, Volume 32, Issue 4, Pages 788-804, ISSN 0147-5967, 10.1016/j.jce.2004.07.001.
- Everett, S.J., Hutchinson, W.K. (2002): The Gravity equation in international economics. Scottish Journal of political economy, vol.49, no.5.
- Feenstra C.R. (1998): Integration of Trade and Disintegration of Production in the Global Economy. The Journal of Economic Perspectives, Vol. 12, No. 4, pp. 31-50.
- Gelman A. (2005): Analysis of variance â why it is more important than ever. Ann. Statist. Volume 33, Number 1, 1-53.
- Head K., Mayer T. (2002): Illusory Border Effects Distance missmeasurement inflates estimates of home bias in trade.
- Kang, Heejoon and Fratianni, Michele U. (2006): International Trade Efficiency, the Gravity Equation, and the Stochastic Frontier-
- Kapoor M., Kelejian H.H., Prucha I.R. (2007): Panel Data Models with Spatially Correlated Error Components. Centre for analytical Finance Research Paper No.02-05.
- Kelejian H.H., Prucha I.R. (1998): a Generalized Spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. Journal of real estate finance and economics, Vol.17:1,99-121.

- LeSage J. and Pace R.K., (2008): Spatial econometrics modeling of origin-destination flows. *Journal of Regional Science*, 48(5): 941-967
- Marti'nez-Zarzoso I., Sua'rez-Burguet C., (2005): Transport costs and trade: Empirical evidence for latino American imports from the European union. *The Journal of International Trade Economic Development*: Vol. 14, Iss. 3.
- Mutl J., Pfaffermayr M. (2008): The Hausman Test in a Cliff and Ord Panel Model. *Econometrics Journal*, volume 10, pp.1-30.
- Politis, Romano (1994): The Stationary Bootstrap. *Journal of the American Statistical Association*, Volume 89, Issue 428.
- Rose, A.K., (2000): One money, one market: Estimating the effect of common currencies on trade. *Economic Policy* 20 (April), 7-45.
- Tinbergen J. (1962): *Shaping the World Economy; Suggestions for an International Economic Policy* - Twenty Century Fund, New York.
- Tobler W.R. (1970): A computer movie simulating urban growth in the Detroit region. *Economic Geography*, vol. 46.